NORTH-HOLLAND

Some Inequalities for Singular Values of Matrix Products*

Bo-Ying Wang and Bo-Yan Xi
Department of Mathematics
Beijing Normal University
Beijing 100875, China

Submitted by George P. H. Styan

ABSTRACT

Let $\sigma_1(C) \geqslant \cdots \geqslant \sigma_n(C)$ denote the singular values of a matrix $C \in \mathbb{C}^{n\times m}$, and let $1 \leqslant i_1 < \cdots < i_k \leqslant n$, $0 < r \in \mathbb{R}$. The main results are $\sum_{t=1}^{k}\sigma_{i_t}^r(AB) \geqslant \sum_{t=1}^{k}\sigma_{i_t}^r(A)\sigma_{n-t+1}^r(B)$ and $\sum_{t=1}^{k}\sigma_t^r(AB) \geqslant \sum_{t=1}^{k}\sigma_{i_t}^r(A)\sigma_{n-i_t+1}^r(B)$, where $A \in \mathbb{C}^{p\times n}$, $B \in \mathbb{C}^{n\times m}$. We also consider the cases for the product of three matrices and more. © 1997 Elsevier Science Inc.

Let $\mathbb{C}^{n\times m}$ denote the vector space of all n by m complex matrices. Denote by $\lambda_1(H) \geqslant \cdots \geqslant \lambda_n(H)$ the eigenvalues of a Hermitian matrix $H \in \mathbb{C}^{n\times n}$, and by $\sigma_1(C) \geqslant \cdots \geqslant \sigma_n(C)$ the singular values of $C \in \mathbb{C}^{n\times m}$, that is, $\sigma_i(C) = \lambda_i^{1/2}(CC^*)$, $i = 1,\ldots,n$. When $C \in \mathbb{C}^{p\times m}$ and $q = \max\{p, m\} < n$, we set $\sigma_{q+1}(C) = \cdots = \sigma_n(C) = 0$. Finally let $i_1,\ldots,i_k$ be integers such that $1 \leqslant i_1 < \cdots < i_k \leqslant n$.

We start with the following known results.

LEMMA 1 [3, 4]. *Let $W_1,\ldots,W_k$, $V_1,\ldots,V_k$ be subspaces of $\mathbb{C}^n$ with $\dim W_t \geqslant i_t$, $\dim V_t \geqslant n - i_t + 1$, $1 \leqslant i_1 < \cdots < i_k \leqslant n$, and $V_1 \supset \cdots \supset V_k$. Then there exists a subspace $W \subset \mathbb{C}^n$ with an orthonormal basis $\{x_1,\ldots,x_k\}$ and another orthonormal basis $\{z_1,\ldots,z_k\}$ such that $x_t \in W_t$ and $z_t \in V_t$, $t = 1,\ldots,k$.*

* Project supported by the National Natural Science Foundation of China.

655 Avenue of the Americas, New York, NY 10010
0024-3795/97/$17.00
PII S0024-3795(97)00020-7

LEMMA 2 (Lidskiĭ [2, p. 248]). *Let $G, H \in \mathbb{C}^{n\times n}$ be positive semidefinite Hermitian, and let $1 \leqslant i_1 < \cdots < i_k \leqslant n$. Then*

$$\prod_{t=1}^{k} \lambda_{i_t}(GH) \leqslant \prod_{t=1}^{k} \lambda_{i_t}(G)\lambda_t(H)$$

and

$$\prod_{t=1}^{k} \lambda_{i_t}(GH) \geqslant \prod_{t=1}^{k} \lambda_{i_t}(G)\lambda_{n-t+1}(H)$$

with equality for $k = n$.

Let $A \in \mathbb{C}^{p\times n}$, $B \in \mathbb{C}^{n\times m}$. Lemma 2 yields

$$\prod_{t=1}^{k} \sigma_{i_t}(AB) = \prod_{t=1}^{k} \lambda_{i_t}^{1/2}(ABB^*A^*) = \left(\prod_{t=1}^{k} \lambda_{i_t}(A^*ABB^*)\right)^{1/2}$$

$$\geqslant \left(\prod_{t=1}^{k} \lambda_{i_t}(A^*A)\lambda_{n-t+1}(BB^*)\right)^{1/2} = \prod_{t=1}^{k} \sigma_{i_t}(A)\sigma_{n-t+1}(B).$$

Thus (Gel'fand and Naimark [2, p. 248]),

$$\prod_{t=1}^{k} \sigma_{i_t}(AB) \geqslant \prod_{t=1}^{k} \sigma_{i_t}(A)\sigma_{n-t+1}(B). \tag{1}$$

LEMMA 3 [5]. *Let $A \in \mathbb{C}^{p\times n}$, $B \in \mathbb{C}^{n\times m}$, and $1 \leqslant i_1 < \cdots < i_k \leqslant n$. Then*

$$\prod_{t=1}^{k} \sigma_t(AB) \geqslant \prod_{t=1}^{k} \sigma_{i_t}(A)\sigma_{n-i_t+1}(B) \tag{2}$$

and

$$\sum_{t=1}^{k} \sigma_t(AB) \geqslant \sum_{t=1}^{k} \sigma_{i_t}(A)\sigma_{n-i_t+1}(B). \tag{3}$$

COROLLARY 1. *Let* $A \in \mathbb{C}^{p\times n}$, $B \in \mathbb{C}^{n\times m}$, *and* $1 \leqslant i_1 < \cdots < i_k \leqslant n$, $0 < r \in \mathbb{R}$. *Then*

$$\prod_{t=1}^{k} \sigma_{i_t}^r(AB) \geqslant \prod_{t=1}^{k} \sigma_{i_t}^r(A)\sigma_{n-t+1}^r(B),$$

$$\prod_{t=1}^{k} \sigma_{t}^r(AB) \geqslant \prod_{t=1}^{k} \sigma_{i_t}^r(A)\sigma_{n-i_t+1}^r(B)$$

$$\sum_{t=1}^{k} \sigma_{t}^r(AB) \geqslant \sum_{t=1}^{k} \sigma_{i_t}^r(A)\sigma_{n-i_t+1}^r(B). \tag{4}$$

Proof. The first two inequalities are obvious by (1) and (2).

The proof of the third inequality, by majorization, is the same as the proof of (3); see [5, Theorem 3]. ∎

Let $u_1, \dots, u_n$ be orthonormal eigenvectors of AA^* corresponding to $\lambda_1(AA^*), \dots, \lambda_n(AA^*)$, where $A \in \mathbb{C}^{n\times m}$, that is,

$$AA^* u_i = \lambda_i(AA^*)u_i = \sigma_i^2(A)u_i, \qquad i = 1, \dots, n.$$

Let $R_t \subset \mathbb{C}^n$ denote the subspace spanned by $\{u_1, u_2, \dots, u_{i_t}\}$, and S_t the subspace spanned by $\{u_{i_t}, u_{i_t+1}, \dots, u_n\}$, $t = 1, \dots, k$. For $x_1, \dots, x_k \in \mathbb{C}^n$, let $X_k = (x_1, \dots, x_k) \in \mathbb{C}^{n\times k}$, $I_k = \operatorname{diag}(1, \dots, 1) \in \mathbb{C}^{k\times k}$.

LEMMA 4. *Let* $A \in \mathbb{C}^{n\times m}$, *and let* $Y_k = (y_1, \dots, y_k)$, $Z_k = (z_1, \dots, z_k) \in \mathbb{C}^{n\times k}$ *with* $y_t \in R_t$, $z_t \in S_t$, $t = 1, \dots, k$.

(i) *If* $Y_k^* Y_k = I_k$, *then* $\sigma_t(Y_k^* A) \geqslant \sigma_{i_t}(A)$, $t = 1, \dots, k$.

(ii) *If* $Z_k^* Z_k = I_k$, *then* $\sigma_t(Z_k^* A) \leqslant \sigma_{i_t}(A)$, $t = 1, \dots, k$.

Proof. For a fixed t $(1 \leqslant t \leqslant k)$, choose $v_{t+1}, \dots, v_{i_t} \in R_t$ such that $\{y_1, \dots, y_t, v_{t+1}, \dots, v_{i_t}\}$ is an orthonormal basis spanning R_t. Then it is easy to check that the matrices $(y_1, \dots, y_t, v_{t+1}, \dots, v_{i_t})^* A$ and $(u_1, \dots, u_{i_t})^* A$ have the same singular values. Applying the intersection property of singular values of a matrix $B \in \mathbb{C}^{q\times m}$ and its submatrix $C \in \mathbb{C}^{p\times m}$, namely $\sigma_i(B) \geqslant$

$\sigma_i(C)$ and $\sigma_{q-i+1}(B) \leqslant \sigma_{p-i+1}(C)$, $i = 1, \dots, p$ (that is, the intersection property of eigenvalues of Hermitian matrix $BB^* \in \mathbb{C}^{q\times q}$ and its principal submatrix $CC^* \in \mathbb{C}^{p\times p}$; see [2, p. 227]), we have

$$\sigma_t(Y_k^*A) \geqslant \sigma_t((y_1, \dots, y_t)^* A) \geqslant \sigma_{i_t}\big((y_1, \dots, y_t, v_{t+1}, \dots, v_{i_t})^* A\big)$$

$$= \sigma_{i_t}\big((u_1, \dots, u_{i_t})^* A\big) = \sigma_{i_t}(A).$$

Similarly,

$$\sigma_t(Z_k^*A) \leqslant \sigma_1((z_t, \dots, z_k)^* A) \leqslant \sigma_1\big((u_{i_t}, \dots, u_n)^* A\big) = \sigma_{i_t}(A). \quad \blacksquare$$

Now we may show the following result.

THEOREM 1. *Let* $A \in \mathbb{C}^{n\times m}$ *and* $1 \leqslant i_1 < \cdots < i_k \leqslant n$. *Let* $\varphi(\alpha_1, \dots, \alpha_k)$ *be any function of* k *variables which is increasing in each variable in the interval* $[0, \infty)$. *Then*

$$\varphi\big(\sigma_{i_1}(A), \dots, \sigma_{i_k}(A)\big) = \max_{\substack{\dim W_t = i_t \\ t=1,\dots,k}} \min_{\substack{x_t \in W_t \\ X_k^* X_k = I_k}} \varphi(\sigma_1(X_k^*A), \dots, \sigma_k(X_k^*A)). \tag{5}$$

Proof. Let W_t be any subspace of $\mathbb{C}^n$ with $\dim W_t = i_t$, $t = 1, \dots, k$. Note that $\dim S_t = n - i_t + 1$, $t = 1, \dots, k$ and $S_1 \supset \cdots \supset S_k$. From Lemma 1, there exists a subspace of $\mathbb{C}^n$ which has two orthonormal bases $\{x_1, \dots, x_k\}$ and $\{z_1, \dots, z_k\}$ such that $x_t \in W_t$ and $z_t \in S_t$, $t = 1, \dots, k$.

Therefore, by Lemma 4(ii) and the fact that φ is an increasing function, we have

$$\varphi(\sigma_1(X_k^*A), \dots, \sigma_k(X_k^*A)) = \varphi(\sigma_1(Z_k^*A), \dots, \sigma_k(Z_k^*A))$$

$$\leqslant \varphi\big(\sigma_{i_1}(A), \dots, \sigma_{i_k}(A)\big).$$

It follows that the left side of (5) is no smaller than its right side.

On the other hand, take $W_t = R_t$, $t = 1, \dots, k$, and suppose $\{y_1, \dots, y_k\}$ is any orthonormal set with $y_t \in R_t$, $t = 1, \dots, k$. Then by Lemma 4(i)

$$\varphi(\sigma_1(Y_k^*A), \dots, \sigma_k(Y_k^*A)) \geqslant \varphi\big(\sigma_{i_1}(A), \dots, \sigma_{i_k}(A)\big).$$

Thus the left side of (5) is no greater than its right side. The proof is completed. ■

This theorem is a generalization of a result due to Amir-Moéz [1, Theorem 2.3], in which there is a restriction $W_1 \subset \cdots \subset W_k$.

THEOREM 2. *Let* $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, *and* $1 \leqslant i_1 < \cdots < i_k \leqslant n$, $0 < r \in \mathbb{R}$. *Then*

$$\sum_{t=1}^{k} \sigma_{i_t}^r(AB) \geqslant \sum_{t=1}^{k} \sigma_{i_t}^r(A)\sigma_{n-t+1}^r(B). \tag{6}$$

Proof. Set $\varphi(\alpha_1,\dots,\alpha_k) = \sum_{t=1}^{k}\alpha_t^r$, and take $W_t = R_t$ in Theorem 1. We obtain by using Corollary 1 and Lemma 4(i)

$$\begin{aligned}\sum_{t=1}^{k} \sigma_{i_t}^r(AB) &\geqslant \min_{\substack{y_t\in R_t\\ Y_k^*Y_k=I_k}} \sum_{t=1}^{k} \sigma_t^r(Y_k^*AB)\\ &\geqslant \min_{\substack{y_t\in R_t\\ Y_k^*Y_k=I_k}} \sum_{t=1}^{k} \sigma_t^r(Y_k^*A)\sigma_{n-t+1}^r(B)\\ &\geqslant \sum_{t=1}^{k} \sigma_{i_t}^r(A)\sigma_{n-t+1}^r(B). \end{aligned}$$ ■

The following is an extension of [5, Theorems 4 and 3], in which $r = 1$.

COROLLARY 2. *Let* $G, H \in \mathbb{C}^{n\times n}$ *be positive semidefinite Hermitian, and let* $1 \leqslant i_1 < \cdots < i_k \leqslant n$, $0 < r \in \mathbb{R}$. *Then*

$$\sum_{t=1}^{k} \lambda_{i_t}^r(GH) \geqslant \sum_{t=1}^{k} \lambda_{i_t}^r(G)\lambda_{n-t+1}^r(H), \tag{7}$$

$$\sum_{t=1}^{k} \lambda_{t}^r(GH) \geqslant \sum_{t=1}^{k} \lambda_{i_t}^r(G)\lambda_{n-i_t+1}^r(H).$$

Proof. The result follows on replacing A, B, and r respectively by $G^{1/2}$, $H^{1/2}$, and $2r$ in (6) and (4). ■

From (7), one can see that Theorem 2 holds for $A \in \mathbb{C}^{p\times n}$, $B \in \mathbb{C}^{n\times m}$ by using the same method as the proof of (1).

Finally, we discuss the cases of the product of m matrices.

THEOREM 3. *Let* $A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$, $m \geqslant 3$, $1 \leqslant i_1 < \cdots < i_k \leqslant n$, $0 < r \in \mathbb{R}$. *Then*

(i) $\prod_{t=1}^k \sigma_{i_t}^r(A_1 \cdots A_m) \leqslant \prod_{t=1}^k \sigma_{i_t}^r(A_1)\sigma_t^r(A_2)\cdots \sigma_t^r(A_m)$;

(ii) $\sum_{t=1}^k \sigma_{i_t}^r(A_1 \cdots A_m) \leqslant \sum_{t=1}^k \sigma_{i_t}^r(A_1)\sigma_t^r(A_2)\cdots \sigma_t^r(A_m)$;

(iii) $\prod_{t=1}^k \sigma_{i_t}^r(A_1 \cdots A_m) \geqslant \prod_{t=1}^k \sigma_{i_t}^r(A_1)\sigma_{n-t+1}^r(A_2)\cdots \sigma_{n-t+1}^r(A_m)$;

(iv) $\prod_{t=1}^k \sigma_t^r(A_1 \cdots A_m) \geqslant \prod_{t=1}^k \sigma_{i_t}^r(A_1)\sigma_{n-i_t+1}^r(A_2)\sigma_{n-t+1}^r(A_3)\cdots \times \sigma_{n-t+1}^r(A_m)$.

Proof. (i) and (ii) are trivial from the known results; see [2, pp. 247, 250].

(iii) and (iv): From Corollary 1, by induction on m, we obtain

$$\begin{aligned}
\prod_{t=1}^k \sigma_{i_t}^r(A_1 \cdots A_m) &\geqslant \prod_{t=1}^k \sigma_{i_t}^r(A_1 \cdots A_{m-1})\sigma_{n-t+1}^r(A_m)\\
&= \left(\prod_{t=1}^k \sigma_{i_t}^r(A_1 \cdots A_{m-1})\right)\prod_{t=1}^k \sigma_{n-t+1}^r(A_m)\\
&\geqslant \left(\prod_{t=1}^k \sigma_{i_t}^r(A_1)\sigma_{n-t+1}^r(A_2)\cdots \sigma_{n-t+1}^r(A_{m-1})\right)\\
&\quad\times \prod_{t=1}^k \sigma_{n-t+1}^k(A_m)\\
&= \prod_{t=1}^k \sigma_{i_t}^r(A_1)\sigma_{n-t+1}^r(A_2)\cdots \sigma_{n-t+1}^r(A_m),
\end{aligned}$$

$$\begin{aligned}
\prod_{t=1}^k \sigma_t^r(A_1 \cdots A_m) &\geqslant \prod_{t=1}^k \sigma_t^r(A_1A_2)\sigma_{n-t+1}^r(A_3)\cdots \sigma_{n-t+1}(A_m)\\
&\geqslant \prod_{t=1}^k \sigma_{i_t}^r(A_1)\sigma_{n-i_t+1}^r(A_2)\sigma_{n-t+1}^r(A_3)\cdots\\
&\quad \sigma_{n-t+1}^r(A_m). \quad ■
\end{aligned}$$

REMARK. For $m \geqslant 3$, because of (6) and (4), it is natural to conjecture that (iii) and (iv) will hold if we replace $\Pi_{t=1}^{k}$ by $\Sigma_{t=1}^{k}$. But that is not true. The following is a counterexample: Let

$$A_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

and take $k = n$, $r = 1$; then

$$\sum_{t=1}^{k} \sigma_t(A_1 A_2 A_3)\ (=9) \ \not\geqslant \ \sum_{t=1}^{k} \sigma_t(A_1)\sigma_{n-t+1}(A_2)\sigma_{n-t+1}(A_3)\ (=13).$$

However, from (6) and (4), we can also obtain some weak inequalities:

COROLLARY 3. *Let* $A_1, \dots, A_m \in \mathbb{C}^{n\times n}$, $m \geqslant 3$, $1 \leqslant i_1 < \cdots < i_k \leqslant n$, $0 < r \in \mathbb{R}$. *Then*

(i) $\Sigma_{t=1}^{k}\sigma_{i_t}^{r}(A_1 \cdots A_m) \geqslant \Sigma_{t=1}^{k}\sigma_{i_t}^{r}(A_1)\sigma_{n-t+1}^{r}(A_2)\sigma_n^{r}(A_3)\cdots \sigma_n^{r}(A_m)$;

(ii) $\Sigma_{t=1}^{k}\sigma_{t}^{r}(A_1 \cdots A_m) \geqslant \Sigma_{t=1}^{k}\sigma_{i_t}^{r}(A_1)\sigma_{n-i_t+1}^{r}(A_2)\sigma_n^{r}(A_3)\cdots \sigma_n^{r}(A_m)$.

Proof. Use (6) and (4), and note that $\sigma_i^{r}(AB) \geqslant \sigma_i^{r}(A)\sigma_n^{r}(B)$ for any $A, B \in \mathbb{C}^{n\times n}$. ■

REFERENCES

1 A. R. Amir-Moéz, Extreme properties of eigenvalues of a Hermitian transformation and singular values of the sum and product of linear transformations, *Duke Math. J.* 23:463–476 (1956).

2 A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic, New York, 1979.

3 B. Y. Wang, On the extreme properties of eigenvalues and the subspace inclusion problem, *Adv. in Math. (China)* 15:431–433 (1986).

4 B. Y. Wang, A conjecture on orthonormal bases, *Linear and Multilinear Algebra* 28:193 (1990).

5 B. Y. Wang and F. Zhang, Some inequalities for the eigenvalues of the product of positive semidefinite Hermitian matrices, *Linear Algebra Appl.* 160:113–118 (1992).

Received 29 January 1996; final manuscript accepted 14 February 1997